

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 June 2001 (21.06.2001)

PCT

(10) International Publication Number
WO 01/45410 A2

(51) International Patent Classification⁷: **H04N 7/16**

(21) International Application Number: **PCT/US00/33151**

(22) International Filing Date: 6 December 2000 (06.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/461,259 15 December 1999 (15.12.1999) US

(71) Applicant: **SUN MICROSYSTEMS, INC. [US/US]**,
901 San Antonio Road, M/S: UPAL01-521, Palo Alto, CA
94303 (US).

(72) Inventors: **CARONNI, Germano**; 1063 Morse Avenue #25-300, Sunnyvale, CA 94089 (US). **SCHUBA, Christoph**; 473 Hope Street #1, Mountain View, CA 94041 (US).

(74) Agents: **HECKER, Gary, A. et al.**; The Hecker Law Group, Suite 2300, 1925 Century Park East, Los Angeles, CA 90067 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

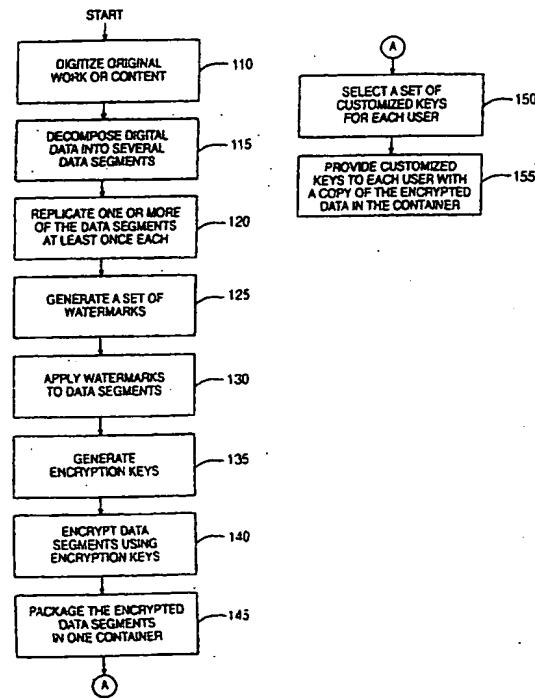
(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— Without international search report and to be republished upon receipt of that report.

[Continued on next page]

(54) Title: A METHOD AND APPARATUS FOR WATERMARKING DIGITAL CONTENT



(57) Abstract: A method and apparatus for watermarking digital data is described herein whereby the digital data is decomposed into a plurality of original data segments, one or more of the original data segments replicated at least once to generate replica data segments, a set of watermarks is generated, each watermark is applied to a respective data segment to generate watermarked data segments, the data segments are encrypted using encryption keys to generate encrypted data segments. One or more embodiments of the invention include providing a subset of the encryption keys corresponding to a subset of the encrypted data segments, wherein each encrypted data segment in the subset of the encrypted data segments, can be decrypted using a corresponding encryption key in the subset of encryption keys, and wherein the decrypted data segments can be combined to reconstruct the digital data including one or more of the watermarks.

WO 01/45410 A2

WO 01/45410 A2



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

A METHOD AND APPARATUS FOR WATERMARKING DIGITAL CONTENT

Field of the Invention

5

The present invention relates generally to preventing illicit copying of digitized works using watermarks, and more particularly, to a mechanism for distribution of works or content for use in the prevention of illicit copying and/or collusion to eliminate the watermarks.

10

Portions of the disclosure of this patent document contain material that is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure as it appears in the Patent and Trademark Office file or records, but 15 otherwise reserves all copyright rights whatsoever. Sun, Sun Microsystems, Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States and other countries. All SPARC trademarks are used under license and are trademarks of SPARC International, Inc. in the United States and other countries. Products bearing 20 SPARC trademarks are based upon an architecture developed by Sun Microsystems, Inc.

Background of the Invention

25

By digitizing information (e.g., audio, text, video, image, computer programs, etc.) one eliminates the problem of quality degradation when a copy of the digitized data may be made and distributed easily without loss in quality.

However, it is possible for digitized information to be copied and distributed without the permission of the owner of the digitized information. Because the copied data is identical for all recipients, it is impossible to determine where and when the first illicit or bootlegged copy was made. As such, there has been an
5 increasing need for methods of tracking copies of digital data distributed to various recipients which cannot be easily defeated.

Watermarking has been employed to address such needs. Watermarking provides for embedding additional, hard to detect watermarks or "tags" into the
10 original digitized data such as images, audio recordings, movies, computer program, etc. For example, portions, or segments, of an image may be modified by shifting the brightness or by displacing contours in a way that there is no perceptible degradation in the appearance or quality of the image. Each legitimate recipient of the image receives a customized, tagged version of the
15 digitized data that is unique to that recipient. If a tagged copy of digitized data is found as a bootlegged copy, it can be analyzed, the watermarks can be identified, and the original recipient may be determined. In so doing, it is possible to determine the party or parties that gave away their customized, legitimate copy with high degree of confidence.

20

A disadvantage of such watermarking techniques is that a digital work must be personalized for each recipient and delivered to each recipient separately. Therefore, bulk delivery of a single digital image by multicasting or CD-ROM distribution, for example, is not feasible. Another disadvantage of
25 such watermarking techniques is that watermarks can often be removed if enough recipients with different legitimate copies collude. It is possible, for example, by comparing enough watermarked versions of the digital work, to

isolate the modifications made to the original work, and recreate the original work, or to interchange enough modifications, such that the resulting version can no longer be attributed to an original, legitimate recipient.

5 There is, therefore, a need for a method of using existing watermarking and encryption techniques to customize digital versions of works to determine the source of illicit copies of such works.

SUMMARY OF THE INVENTION

Embodiments of the invention comprise a method and apparatus for watermarking digital data wherein a digitized version of a work is decomposed
5 into a plurality of original data segments. The plurality of original data segments are replicated at least once to generate replica data segments. A set of watermarks (i.e., modifications to the digital information) are identified for at least one of the original data segments. A watermarked data segment is generated by applying a watermark to a replica or original data segment. That
10 is, for example, a copy can be made of the original data prior to the application of watermark(s). In an alternate embodiment, a copy is generated in the course of applying watermark(s) to the original data.

A set of encryption keys is generated and a data segment (e.g., an original or replica data segment with or without a watermark) is encrypted using a
15 respective encryption key. A distribution set is generated comprising encrypted data segments wherein some or all of the encrypted data segments contain a watermark. In one or more embodiments of the invention, each data segment further contains location information that may be used in combining data segments to reconstruct the work. In one or more embodiments of the
20 invention, a set of encryption keys is distributed to a recipient to decrypt one or more data segments that comprise a unique version of the work including one or more watermarks. The encrypted data segments can be decrypted using the encryption keys to recover the unique version of the work.

25 Using to one or more embodiments of the invention, a unique version of the work may be given to each recipient wherein each unique version comprises

a unique combination of watermarks. Since each recipient is associated with a unique combination of watermarks, embodiments of the invention aid in determining the source of an illicit copy. According to one or more embodiments of the invention, an illicit copy may be compared with the original 5 digital data to identify the unique combination of watermarks contained in the copy. The unique combination of watermarks may be used to identify the copy's original recipient.

According to one or more embodiments of the invention, the distribution set comprises all of the data segments (e.g., original and replica) each encrypted 10 using their respective encryption keys. The distribution set may be distributed on a CD-ROM or multicasted, for example, to all recipients. Each recipient is provided with a set of keys that may be used to decrypt an associated set of the data segments that comprise a personalized version of the original work. Thus, embodiments of the invention facilitate a bulk delivery of personalized content.

15 Instead of distributing all data segments, in one or more embodiments of the invention, a recipient is provided with one or more keys and an identifier associated with each key and a data segment. To retrieve the personalized set of data segments, the recipient provides the set of keys which are used to determine the set of data segments that are to be delivered to the recipient.

20 Those data segments that are associated with the identifiers provided may be delivered electronically (e.g., via the Internet) to the recipient thereby reducing the amount of data that is to be transmitted to each recipient.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1B illustrate watermarking process flows according to one or more embodiments of the invention.

5

Figures 2A-2B illustrate process flows wherein watermarked digital data is reconstructed according to one or more embodiments of the invention.

Figures 3A-3B provide examples of watermarking digital data according to one or more embodiments of the invention.

10 Figure 4 illustrates a process flow of identifying the source of an illicit (or unauthorized) copy of watermarked digital data according to one or more embodiments of the invention.

15 Figure 5 is a block diagram of one embodiment of a computer system capable of providing a suitable execution environment for an embodiment of the invention.

Figure 6A-6B provide a block-level overview according to one or more embodiments of the invention.

Figure 7 provides a block-level watermarking overview using parallelism according to one or more embodiments of the invention.

DETAILED DESCRIPTION OF THE INVENTION

A method and apparatus for watermarking digital content is described. In the following description, numerous specific details are set forth in order to provide a more thorough description of the present invention. It will be 5 apparent, however, to one skilled in the art, that the present invention may be practiced without these specific details. In other instances, well-known features have not been described in detail so as not to obscure the invention.

The present invention provides for watermarking of digitized 10 information such as digitized text, image, audio, video, program code, etc. The present invention also provides for access control or rights management via encryption, where different encryption keys are used to encrypt, for example, different segments of digital data to make the digital data accessible to selected parties only.

15 Figures 1A-1B illustrate watermarking process flows according to one or more embodiments of the invention. At step 110, an original work is encoded as digital data. For example, an original digitized image can be decomposed into a patchwork of possibly overlapping data segments representing rectangles. The rectangles are then replicated and watermarked. Thus, at step 115, the digital 20 data is decomposed into a plurality of original data segments and the data segments' respective locations within the original work is determined. One or more of the original data segments are replicated at least once to generate replica data segments at step 120.

According to an embodiment of the invention, replicating one or more of the original data segments can include selecting a subset of the original data segments, and replicating each data segment in the selected subset at least once to generate said replica data segments. Alternatively, each original data segment
5 can be replicated at least once to generate the replica data segments.

At step 125, a watermark is identified (e.g., an area of the digitized data is identified along with the modifications to be made to the area). In one example for watermarking a digital image according to one or more embodiments of the
10 invention, locations on the original digital image where watermarks can be applied are identified. Then the original digital image is watermarked differently for each intended recipient of the digital image. For example, watermarking can include shifting the brightness in portions of the digital image, or displacing image contours. A watermarked image can include many hundred or several
15 thousand locations (either in the space or in the frequency domain) where watermarking, in the form of minor changes to the image, is applied. The process of watermarking can be automated for efficiency.

A watermark is applied to a respective data segment to generate a watermarked data segment at step 130. Steps 125 and 130 may be repeated any
20 number of times to generate additional watermarked data segments. In an embodiment of the invention, each original and each replica data segment is watermarked. In alternate embodiments of the invention, some number (e.g., less than all) of the original and replica data segments are watermarked.

According to one or more embodiments of the invention, each data segment can
25 be replicated several times, and depending on the number of replica data segments and the number of data segments watermarked, several watermarked

versions of the original work can be generated by choosing different combinations of watermarked data segments.

In the example of Figures 1A-1B, watermarks may be applied to a copy of the original data. Other techniques may be used, however. For example, in an
5 alternate embodiment, a watermark may be applied to an original data segment to produce a watermarked replication of the original data segment.

Each data segment, along with its location information (e.g., image coordinates), is encrypted and augmented by a respective encryption key (e.g., a symmetric encryption key). That is, a set of encryption keys are generated at
10 step 135. Each data segment (e.g., each original and replica data segment) is encrypted with a key to generate encrypted data segments at step 140. For example, a unique key can be generated for each data segment and the data segments can be encrypted with different cryptographic symmetric keys. Using symmetric keys, the same key is used for both encryption and decryption of a
15 data segment.

While embodiments of the invention are illustrated with reference to the use of symmetric keys, it is possible to use other encryption schemes (e.g., asymmetric key encryption wherein different keys are used for encryption and decryption) with one or more embodiments of the invention. In one or more
20 embodiments of the invention, one key is used to encrypt and decrypt one data segment. Alternatively, the same key may be used to encrypt (and decrypt) two or more data segments.

In one or more embodiments of the invention, the encrypted data segments are then packaged in a single container (or distribution set) or media
25 for bulk distribution to different recipients at step 145. As such, instead of

delivering a personalized copy of the digital data for every recipient, all the recipients receive a copy of the same encrypted data segments. According to one or more embodiments of the invention, the distribution set contains multiple copies (in data segments) of the original digital data with watermarks 5 and location (or reconstruction) information. According to one or more embodiments of the invention, a recipient can recover the coordinate information for each rectangle in the image only after decryption.

A set of customized keys are selected for each recipient at step 150 where the customized keys allow the recipient to decrypt enough of the encrypted data 10 segments to reconstruct one version of the original work with watermarks therein. Each customized set of keys can include a unique subset of the encryption keys used to encrypt all the data segments. Each recipient is provided with a copy of the encrypted data segments and customized keys for that recipient at step 155.

15 In the example of Figure 1A, steps 150 and 155 are not necessarily performed automatically after a container is formed in step 145. For example, steps 150 and 155 may be invoked in response to a user request, payment or other stimulus. If, for example, in a user interaction (or user interaction phase), a user provides payment, steps 150 and 155 may be performed in response to the 20 user's payment.

The process flow of Figure 1A may be modified such that portions of the flow may be performed in parallel. Figure 1B provides a watermarking process flow using parallel processing of steps 110-145 (e.g., a preparation phase) according to one or more embodiments of the invention. The parallel 25 processing may be performed using multiple computer systems or multiple processors in a multiprocessor computing systems.

Referring to Figure 1B, for example, data segment generation 170 (e.g., steps 110, 115 and 120) may be performed at the same time as watermark generation 172 (e.g., step 125) and encryption key generation 174 (e.g., step 135).

5 Similarly, watermarking 176 (e.g., steps 110, 115, 120, 125 and 130) may be performed at the same time as encryption key generation 174.

The output of watermarking 176 and encryption key generation 174 become input to step 140 wherein the watermarked data segments obtained in 10 watermarking 176 are encrypted using the encryption keys generated in encryption key generation 174. At step 145, the encrypted data segments are packaged in a container.

According to one or more embodiments of the invention, with a copy of 15 the encrypted data and a set of customized keys, a recipient may decrypt the data sets encrypted using the keys to reconstruct a customized version of the original work that contains a unique set of watermarks. Figures 2A-2B illustrate process flows wherein watermarked digital data is reconstructed according to one or more embodiments of the invention. Upon receiving the encrypted data 20 at step 260 and customized keys at step 265, a recipient utilizes the customized keys to decrypt a subset of the encrypted data segments to recover watermarked data segments at step 270. The decrypted data segments are used to reconstruct a customized version of the original work with watermarks therein at step 275. In one or more embodiments of the invention, the location 25 information associated with each data segment is used to piece the data segments together to reconstruct the work including watermarks.

Referring to Figure 2B, the process flow of Figure 1A is restructured to illustrate use of parallel processing that may be used in one or more embodiments of the invention. The receipt of encrypted data segments (e.g., step 260) may be performed at the same time as receipt of customized keys (e.g., 5 step 265). The encrypted data segments and encryption keys are used (e.g., in step 270) to decrypt a subset of the encrypted data segments. The decrypted data segments are combined (e.g., at step 275) to form a watermarked version of the original data segment.

10 In one or more embodiments of the invention, the same encrypted information is packaged as a distribution set that may be fixed in a type of media (e.g., CD-ROM, removable disk drives, etc.), or multicasted such that all or a portion of the distribution set, and delivered to a recipient along with a customized set of keys for that recipient. Different sets of keys allow recipients 15 to decrypt different image rectangles, leading to different watermarked images for specific combinations of keys provided to the recipients.

To recover an image from the distribution set, a recipient uses the set of keys to decrypt a subset of the original rectangles, wherein the rectangles in the 20 subset together form one instance of the image that is known to have been given to that recipient. The recovered image is a uniquely watermarked instance of the original image.

In one or more embodiments of the invention, the distribution set 25 comprises all of the data segments (i.e., original and replica data segments) numbering $N \times (M+1)$ where N is the number of original data segments and M is the number of replications of the original data segments. For example, referring

to Figure 3A, original image 302 (e.g., an original digital version of an image or a copy of an original digital version) is decomposed into four (4) rectangles, or original data segments, and each rectangle is replicated two (2) times. In the example of Figure 3A, the data segments are non-overlapping. It should be
5 appreciated, however, that one or more data segments may overlap with each other.

The four (4) original rectangular data segments are replicated twice making a total of $4 \times (2+1)$ or twelve (12) data segments (i.e., 4 original data segments and 4 data segments in each replication of the original data segments).
10 Some or all of the $N \times (M+1)$ rectangles are watermarked, where one watermark for each rectangle can be unique. Referring to Figure 3A, watermarks 306 are applied to data segments 304 such that each of watermarks 306 is applied to one of data segments 304 to yield watermarked data segments 308. Watermarks 306 may be different or the same watermark according to one or more
15 embodiments of the invention. Further, different watermarks may be applied to the same data segment.

According to one or more embodiments of the invention, location information in the form of coordinates are associated with each data segment (e.g., the "x,y" coordinate of the top left-hand corner of a rectangle in relation the
20 top left-hand corner of the whole image). Referring to Figure 3A, for example, each of watermarked data segments 308 is encrypted along with their respective location information using a respective encryption key from encryption keys 310 which results in encrypted data segments 312.

According to one or more embodiments of the invention, the entire set of
25 encrypted $N \times (M+1)$ data segments (e.g., data segments 312) are provided to

each recipient in a distribution set (e.g., a distribution set comprising encrypted data segments 312). Each recipient receives a different subset of the encryption keys (e.g., encryption keys 310) for decrypting a different subset of the encrypted Nx(M+1) rectangles to form a different version of the original image 5 including watermarks. For example, in the example of Figure 3A, one recipient may receive encryption keys to decrypt the WA1, WB2, WC1 and WD3 data segments while another recipient may receive encryption keys to decrypt the WA2, WB2, WC2 and WD3 data segments. Each recipient receives one or more watermarked and encrypted replicas for each original data segment.

10

In the example of Figure 3A, there are an equal number of copies of data segments A-D of original image 302. It should be apparent, however, that there may be a different number of replications of data segments A-D. Referring to 15 Figure 3B, for example, data segments A and C are replicated two (2) times while data segments B and D are replicated once and seven (7) times, respectively, as illustrated in data segments 324.

Watermarks 326 are applied to data segments 324 to yield watermarked 20 data segments 328. Watermarks 326 may be different or the same watermark according to one or more embodiments of the invention. Further, different watermarks may be applied to the same data segment. Encryption keys 330 are used to encrypt location information associated with watermarked data segments 328 and watermarked data segments 328 to yield encrypted data 25 segments 332.

According to an embodiment of the invention, encrypted, watermarked data segments 332 comprise a distribution set that includes a number of data segments that can be determined using the following:

5 N
 $\sum_{i=1}^{M_i} + 1$

where "N" is the number of original data segments, "i" represents an index of the data segments, and "M" is the number of replications for a given data segments.

10 Thus, for example, the number of data segments that comprise encrypted data segments 332 is:

$$(M_A+1)+(M_B+1)+(M_C+1)+(M_D+1), \text{ or}$$
$$(2+1)+(1+1)+(2+1)+(7+1)=16$$

15 Embodiments of the invention are described herein with reference to image data. However, embodiments of the invention may be applied to other forms of digital information including text, audio recordings, motion pictures, computer programs, etc. Different techniques of applying the watermarks may be used depending on the type of original work. For example, as described above, where the original work is an image, the watermarks can comprise changes in the image attributes. For example, text, audio and motion picture works may include subtle modifications such as undetectable visual or audible modifications.

Where the work is a computer program, the watermarks can comprise different implementations of each functional module of the computer program. For distribution of a computer program, for example, the program code can be decomposed into several data segments representing different functional 5 modules. For each of the functional modules there are a number of ways to implement the functionality of that module and therefore encode watermark information through implementation choices. A change in functionality may comprise a change in the data and/or behavior, for example.

10 Therefore, each functional module can be implemented in multiple forms to provide similar functionality, wherein each implementation can serve as a watermarked version of the functional module. The different implementations are encrypted with different keys and placed in a container for distribution to various recipients with a unique subset of the keys for each recipient to 15 reconstruct the program code, in a manner similar to that described above.

Other techniques of generating watermarks are known to the practitioners in the art and contemplated by the present invention. Preferably, a unique watermark is generated for each original and replica data segment in the 20 above steps. Alternatively, one watermark can be used for two or more data segments. The watermarks can further include identification information.

Embodiments of the invention may be used to facilitate bulk distribution of multiple, customized versions of digital data that are retrievable by the 25 recipient using the distribution set. The customized version given to a recipient is known such that it is possible to determine, given a customized version of the

digital data, the identity of the original recipient. If an unauthorized party is in possession of a recipient's customized version of the original data, the unique combination of watermarks may be used to identify the source (i.e., the recipient) of the customized version.

5

Figure 4 illustrates a process flow of identifying the source of an illicit (or unauthorized) copy of watermarked digital data according to one or more embodiments of the invention. At step 480, an illicit copy is obtained. To determine the source of an illicit copy of the watermarked data, one or more 10 watermarks from the illicit copy are detected step 485. For example, segments of the copy are compared with the original, non-watermarked data segments to identify the one or more watermarks contained in the copy. The one or more watermarks are utilized to determine the identity of a recipient to whom a legitimate copy with said watermark was provided at step 490.

15

The present invention can also be used against multiple collaborators or where one person has multiple illicit copies of digital data. The scheme can be used to trace identity to at least one user if the number of collaborators or copies does not exceed the log of the number of blocks used in the data. For example, 20 if there are 2 to the 10th blocks, the system would provide the ability to identify at least one user if up to 10 collaborators or copies were used.

As described above, embodiments of the invention are contemplated for use in offline bulk distribution of digital data. Alternatively, online, (e.g., 25 Internet) distribution of a customized version of the digital data and or keys is also contemplated wherein only a recipient's data segments are distributed online.

In an online distribution, it is advantageous to minimize the amount of data transmitted to a recipient. A distribution set that contains multiple copies of the original data segments with watermarks and location information may be large, however. For example, for an image including 1.6 megabytes (MB) of data with 2000 8x8 locations for watermarking, an additional 100-200 kilobytes (KB) of storage space is required for watermarks leading to an overhead of 8-16% per copy distributed.

Further, each recipient receives a set of keys. For example, where there are 2,000 data segments encrypted using a unique key, each recipient has at least 2,000 keys requiring approximately 50 KB of storage space. This adds approximately 4% more overhead per recipient in the example above. Advantageously, however, online distribution used in embodiments of the invention reduces the amount of data transmission by sending only those encrypted data segments that comprise a recipient's customized version. This reduces the amount of data volume to be transmitted online over the amount transmitted with offline distribution of the content where space is typically not an issue.

In one or more embodiments of the invention, online delivery of content (e.g., an image) is performed by transmitting (e.g., via the Internet), a set of encryption keys and associated identifiers to a recipient. The recipient can inform a delivery server of the identifiers associated with the keys previously provided to the recipient without informing the server of the actual keys. The server then preselects from among the encrypted rectangles and transmits to the recipient a subset of the encrypted rectangles encrypted using the keys associated with the identifiers supplied by the recipient. The recipient may

proceed to decrypt the data segment using the recipient's keys. Thus, there is a savings in the bandwidth overhead that would otherwise be necessary if all the encrypted rectangles were transmitted to the recipient.

Figures 6A-6B provide a block-level overview according to one or more embodiments of the invention. Referring to Figure 6A, digital data 602 is input to data segment generator 604 to generate data segments 606 of digital data 602. As described above, the data segments may be overlapping. According to one or more embodiments of the invention, data segment generator 604 generates location information for each of data segments 606. Data segments 606 are input to watermark generator 608 that identifies watermarks and applies a watermark to some or all of data segments 606. The resulting data segments and their location information are represented as data segments 610 which become input to encryption module 614 alone with encryption keys 612. One of encryption keys 612 is applied to one of data segments 610 to generate encrypted data segments that comprise distribution 616.

In an offline distribution approach of one or more embodiments of the invention, distribution set 616 stored on a distributable medium (e.g., CD-ROM) that is distributed to the recipients. Referring to Figure 6B, each recipient also receives recipient keys 622 which along with distribution set 614 become input to decryption module 624. Decryption module 624 decrypts the data segments associated with recipient keys 622 to generate decrypted data segments 626. Decrypted data segments 626 (including watermarks) along with their respective location information become input to reconstructor 628 to generate customized copy 630.

Figure 7 provides a block-level watermarking overview using parallelism according to one or more embodiments of the invention. Digital data 722 is input to data segment generator 702 to generate data segments 724 of digital data 722. None or more of data segments 724 may be overlapping with one or 5 more data segments. According to one or more embodiments of the invention, data segment generator 702 generates location information which is included in each of data segments 724.

Watermark generator 704 and key generator 706 may run in parallel with data segment generator 702 to generate watermarks 726 and encryption keys 10 728, respectively. Key generator 706 may further run in parallel with watermark applicator 708. Watermark applicator 708 generates watermarked data segments 730 using data segments 724 and watermarks 726 as input.

Watermarked data segments 730 are input to encryptor 710 along with encryption keys 728. One of encryption keys 728 is applied to one of data 15 segments 730 to generate encrypted data segments that comprise distribution 732.

A detection mechanism may be used in one or more embodiments of the invention to detect an illicit copy. The detection mechanism, or detector, is configured to detect a number of watermarks in data segments that comprise a 20 copy of digital data. Each recipient receives a unique combination of watermarks. Therefore, using a unique combination of watermarks in the copy, it is possible to determine the original recipient of the copy. If the current copy holder is not the original recipient, or is not an authorized recipient, the copy may be considered an illicit copy, for example.

Embodiment of Computer Execution Environment (Hardware)

An embodiment of the invention can be implemented as computer software in the form of computer readable code executed on a general purpose computer such as computer 500 illustrated in Figure 5, or in the form of 5 bytecode class files executable within a runtime environment (e.g., a Java runtime environment) running on such a computer. A keyboard 510 and mouse 511 are coupled to a bi-directional system bus 518. The keyboard and mouse are for introducing user input to the computer system and communicating that user input to processor 513. Other suitable input devices may be used in addition to, 10 or in place of, the mouse 511 and keyboard 510. I/O (input/output) unit 519 coupled to bi-directional system bus 518 represents such I/O elements as a printer, A/V (audio/video) I/O, etc.

Computer 500 includes a video memory 514, main memory 515 and mass storage 512, all coupled to bi-directional system bus 518 along with keyboard 15 510, mouse 511 and processor 513. The mass storage 512 may include both fixed and removable media, such as magnetic, optical or magnetic optical storage systems or any other available mass storage technology. Bus 518 may contain, for example, thirty-two address lines for addressing video memory 514 or main memory 515. The system bus 518 also includes, for example, a 32-bit data bus 20 for transferring data between and among the components, such as processor 513, main memory 515, video memory 514 and mass storage 512. Alternatively, multiplex data/address lines may be used instead of separate data and address lines.

In one embodiment of the invention, the processor 513 is a 25 microprocessor manufactured by Motorola, such as the 680X0 processor or a

microprocessor manufactured by Intel, such as the 80X86, or Pentium processor, or a SPARC microprocessor from Sun Microsystems, Inc. However, any other suitable microprocessor or microcomputer may be utilized. Main memory 515 is comprised of dynamic random access memory (DRAM). Video memory 514 is a
5 dual-ported video random access memory. One port of the video memory 514 is coupled to video amplifier 516. The video amplifier 516 is used to drive the cathode ray tube (CRT) raster monitor 517. Video amplifier 516 is well known in the art and may be implemented by any suitable apparatus. This circuitry converts pixel data stored in video memory 514 to a raster signal suitable for use
10 by monitor 517. Monitor 517 is a type of monitor suitable for displaying graphic images. Alternatively, the video memory could be used to drive a flat panel or liquid crystal display (LCD), or any other suitable data presentation device.

Computer 500 may also include a communication interface 520 coupled to bus 518. Communication interface 520 provides a two-way data communication
15 coupling via a network link 521 to a local network 522. For example, if communication interface 520 is an integrated services digital network (ISDN) card or a modem, communication interface 520 provides a data communication connection to the corresponding type of telephone line, which comprises part of network link 521. If communication interface 520 is a local area network (LAN)
20 card, communication interface 520 provides a data communication connection via network link 521 to a compatible LAN. Communication interface 520 could also be a cable modem or wireless interface. In any such implementation, communication interface 520 sends and receives electrical, electromagnetic or optical signals which carry digital data streams representing various types of
25 information.

Network link 521 typically provides data communication through one or more networks to other data devices. For example, network link 521 may provide a connection through local network 522 to local server computer 523 or to data equipment operated by an Internet Service Provider (ISP) 524. ISP 524 in 5 turn provides data communication services through the world wide packet data communication network now commonly referred to as the "Internet" 525. Local network 522 and Internet 525 both use electrical, electromagnetic or optical signals which carry digital data streams. The signals through the various networks and the signals on network link 521 and through communication 10 interface 520, which carry the digital data to and from computer 500, are exemplary forms of carrier waves transporting the information.

Computer 500 can send messages and receive data, including program code, through the network(s), network link 521, and communication interface 520. In the Internet example, remote server computer 526 might transmit a 15 requested code for an application program through Internet 525, ISP 524, local network 522 and communication interface 520.

The received code may be executed by processor 513 as it is received, and/or stored in mass storage 512, or other non-volatile storage for later execution. In this manner, computer 500 may obtain application code in the 20 form of a carrier wave.

Application code may be embodied in any form of computer program product. A computer program product comprises a medium configured to store or transport computer readable code or data, or in which computer readable code or data may be embedded. Some examples of computer program products 25 are CD-ROM disks, ROM cards, floppy disks, magnetic tapes, computer hard

drives, servers on a network, and carrier waves.

The computer systems described above are for purposes of example only. An embodiment of the invention may be implemented in any type of computer system or programming or processing environment, including embedded devices (e.g., web phones, etc.) and "thin" client processing environments (e.g., network computers (NC's), etc.) that support a virtual machine.

Thus, a method and apparatus for watermarking digital content has been described in conjunction with one or more specific embodiments. The invention is defined by the claims and their full scope of equivalents.

CLAIMS

What is claimed is:

1. A method of watermarking digital data, comprising the steps of:
 - 5 obtaining a plurality of data segments from digital data;watermarking a number of said plurality of data segments, said number of said plurality of data segments comprising replicas of said plurality of data segments;
 - 10 encrypting each of said plurality of data segments using one of a plurality of encryption keys to generate a plurality of encrypted data segments.
2. The method of claim 1 further comprising selecting a subset of encryption keys from said plurality of encryption keys corresponding to a subset of encrypted data segments from said plurality of encrypted data.
3. The method of claim 2 further comprising:
 - 15 decrypting each data segment in said subset of encrypted data segments using a corresponding encryption key in said subset of encryption keys, and combining the decrypted data segments to reconstruct said digital data including one or more of said watermarks.
4. The method of claim 1 further comprising:
 - 20 detecting at least one watermark in an illicit copy of the watermarked digital data to determine a source of copying.

5. The method of claim 4 wherein said detecting at least one watermark further comprises:

comparing said illicit copy with an original to detect said one or more watermarks in said illicit copy.

5 6. The method of claim 1 wherein said watermarking a number of said plurality of data segments further comprises:

selecting a subset of said plurality of data segments;

replicating each of said plurality of data segments in the selected subset at least once; and

10 applying a watermark to each of said replica data segments.

7. The method of claim 1 wherein said watermarking a number of said plurality of data segments further comprises:

applying a unique watermark to said number of said plurality of data segments.

15 8. The method of claim 1 wherein said encrypting each of said plurality of data segments further comprises:

generating a unique encryption key for each of said plurality of data segments.

9. The method of claim 1 wherein said digital data comprises at least
20 one image.

10. The method of claim 1 wherein said digital data comprises audio signals.

11. The method of claim 1 wherein said digital data comprises video signals.

12. The method of claim 1 wherein said digital data comprises software program code.

5 13. A method of providing digital data comprising data segments, said data segments comprising a plurality of original data segments and replicas of said original data segments, each of said data segments being encrypted using a respective one of a plurality of encryption keys, wherein a plurality of said data segments include a watermark, said method comprising:

10 providing a subset of the encryption keys corresponding to a subset of said data segments, wherein each data segment in said subset of data segments can be decrypted using a corresponding encryption key in said subset of encryption keys;

15 combining said decrypted data segments to reconstruct said digital data including one or more of said watermarks.

14. The method of claim 13 further comprising:

providing said data segments and said subset of encryption keys on a storage medium.

15. The method of claim 13 further comprising:

20 providing said data segments and said subset of encryption keys over a transmission line.

16. The method of claim 13 further comprising:

providing said data segments and said subset of encryption keys over a network.

17. The method of claim 14 further comprising:

5 applying a unique watermark to said at least one original data segment and said replicas of said at least one original data segment.

18. The method of claim 13 wherein each key in said plurality of encryption keys is unique.

19. The method of claim 13 wherein said digital data comprises at least 10 one image.

20. The method of claim 13 wherein said digital data comprises audio signals.

21. The method of claim 13 wherein said digital data comprises video signals.

15 22. The method of claim 13 wherein said digital data comprises software program code.

23. The method of claim 13 further comprising:

detecting one or more watermarks in an illicit copy of the digital data to determine a source of copying.

24. The method of claim 23 wherein said detecting one or more watermarks further comprises:

comparing said illicit copy with an original of said digital data to detect said one or more watermarks.

5 25. The method of claim 13 wherein said combining said decrypted data segments to reconstruct said digital data further comprises:

decrypting each data segment in said subset of data segments using a corresponding encryption key in said subset of encryption keys; and

combining said decrypted data segments to reconstruct said digital data
10 including one or more watermarks.

26. A computer program product comprising:

a computer usable medium having computer readable program code embodied therein configured to watermark digital data comprising:
computer readable program code configured to cause a computer to

15 obtain a plurality of data segments from digital data;

computer readable program code configured to cause a computer to watermark a number of said plurality of data segments, said number of said plurality of data segments comprising replicas of said plurality of data segments;

20 computer readable program code configured to cause a computer to encrypt each of said plurality of data segments using one of a plurality of encryption keys to generate a plurality of encrypted data segments.

27. The computer program code of claim 26 further comprising:
computer readable program code configured to cause a computer to
select a subset of encryption keys from said plurality of encryption keys
corresponding to a subset of encrypted data segments from said plurality of
5 encrypted data segments.

28. The computer program product of claim 27 further comprising:
computer readable program code configured to cause a computer to
decrypt each data segment in said subset of encrypted data segments using a
corresponding encryption key in said subset of encryption keys; and
10 computer readable program code configured to cause a computer to
combine said decrypted data segments to reconstruct said digital data including
one or more of said watermarks.

29. The computer program product of claim 26 further comprising
computer readable program code configured to cause a computer to detect at
15 least one watermark in an illicit copy of the watermarked digital data to
determine a source of copying.

30. The computer program product of claim 29 wherein said computer
readable program code configured to cause a computer to detect one or more
watermarks further comprises:
20 computer readable program code configured to cause a computer to
compare said illicit copy with an original to detect said one or more watermarks
in said illicit copy.

31. The computer program product of claim 26 wherein said computer readable program code configured to cause a computer to watermark a number of said plurality of data segments further comprises:

- computer readable program code configured to cause a computer to
- 5 select a subset of said plurality of data segments;
- computer readable program code configured to cause a computer to replicate each of said plurality of data segments in the selected subset at least once; and
- computer readable program code configured to cause a computer to
- 10 apply a watermark to each of said replica data segments.

32. The computer program product of claim 26 wherein said computer readable program code configured to cause a computer to watermark a number of said plurality of data segments further comprises:

- computer readable program code configured to cause a computer to
- 15 apply a unique watermark to said number of said plurality of data segments.

33. The computer program product of claim 26 wherein said computer readable program code configured to cause a computer to encrypt each of said plurality of data segments further comprises:

- computer readable program code configured to cause a computer to
- 20 generate a unique encryption key for each of said plurality of data segments.

34. The computer program product of claim 26 wherein said digital data comprises at least one image.

35. The computer program product of claim 26 wherein said digital data comprises audio signals.

36. The computer program product of claim 26 wherein said digital data comprises video signals.

5 37. The computer program product of claim 26 wherein said digital data comprises software program code.

38. A computer program product comprising:

a computer usable medium having a computer readable program code embodied therein configured to provide digital data comprising data segments, 10 said data segments comprising a plurality of original data segments and replicas of said original data segments, each of said data segments being encrypted using a respective one of a plurality of encryption keys, wherein a plurality of said data segments include a watermark, said computer program product comprising:

computer readable program code configured to cause a computer to 15 provide a subset of said plurality of encryption keys corresponding to a subset of said plurality of data segments, wherein each data segment in said subset of data segments can be decrypted using a corresponding encryption key in said subset of encryption keys;

computer readable program code configured to cause a computer to 20 combine said decrypted data segments to reconstruct said digital data including at least one watermark.

39. The computer program product of claim 38 further comprising:

computer readable program code configured to cause a computer to provide said data segments and said subset of encryption keys on a storage medium.

5 40. The computer program product of claim 38 further comprising:

computer readable program code configured to cause a computer to provide said data segments and said subset of encryption keys over a transmission line.

41. The computer program product of claim 38 further comprising:

10 computer readable program code configured to cause a computer to provide said data segments and said subset of encryption keys over a network.

42. The computer program product of claim 38 further comprising:

computer readable program code configured to cause a computer to apply a unique watermark to said at least one original data segment and said 15 replicas of said at least one original data segment.

43. The computer program product of claim 38 wherein each key in said plurality of encryption keys is unique.

44. The computer program product of claim 38 wherein said computer readable program code configured to cause a computer to combine said decrypted data segments to reconstruct said digital data further comprises:

- computer readable program code configured to cause a computer to
- 5 decrypt each data segment in said subset of data segments using a corresponding encryption key in said subset of encryption keys; and
- computer readable program code configured to cause a computer to combine said decrypted data segments to reconstruct said digital data including one or more watermarks.

10 45. A watermarking system comprising:

- data segment generator configured to generate data segments from digital data;
- watermarker coupled to said data segment generator, said watermark generator configured to watermark a number of said data segments;
- 15 encryptor coupled to said watermark generator, said encryptor configured to encrypt said data segments using a plurality of encryption keys.

46. The system of claim 45 wherein said watermark further comprises:

- watermark generator configured to generate a plurality of watermarks
- 20 for at least one of said data segments; and
- watermark applicator coupled to said watermark generator and said data segment generator, said watermark applicator configured to apply said plurality of watermarks to said at least one of said data segments.

47. The system of claim 45 further comprising:
decryptor configured to decrypt a subset of said encrypted data segments
using a subset of said plurality of encryption;
reconstructor coupled to said decryptor, said reconstructor configured to
5 combine said decrypted data segments to reconstruct a copy of said digital data,
a copy of said digital data comprising a subset of said data segments wherein
said subset of said data segments having a unique combination of watermarks.

48. The system of claim 45 wherein said data segments comprise a
plurality of original data segments and a plurality of replicas of each of said
10 plurality of original data segments.

49. The system of claim 45 further comprising:
detector configured to detect at least one watermark in a copy of said
digital data.

50. The system of claim 49 wherein said detector is configured to detect
15 said at least one watermark in said copy to determine a source of copying.

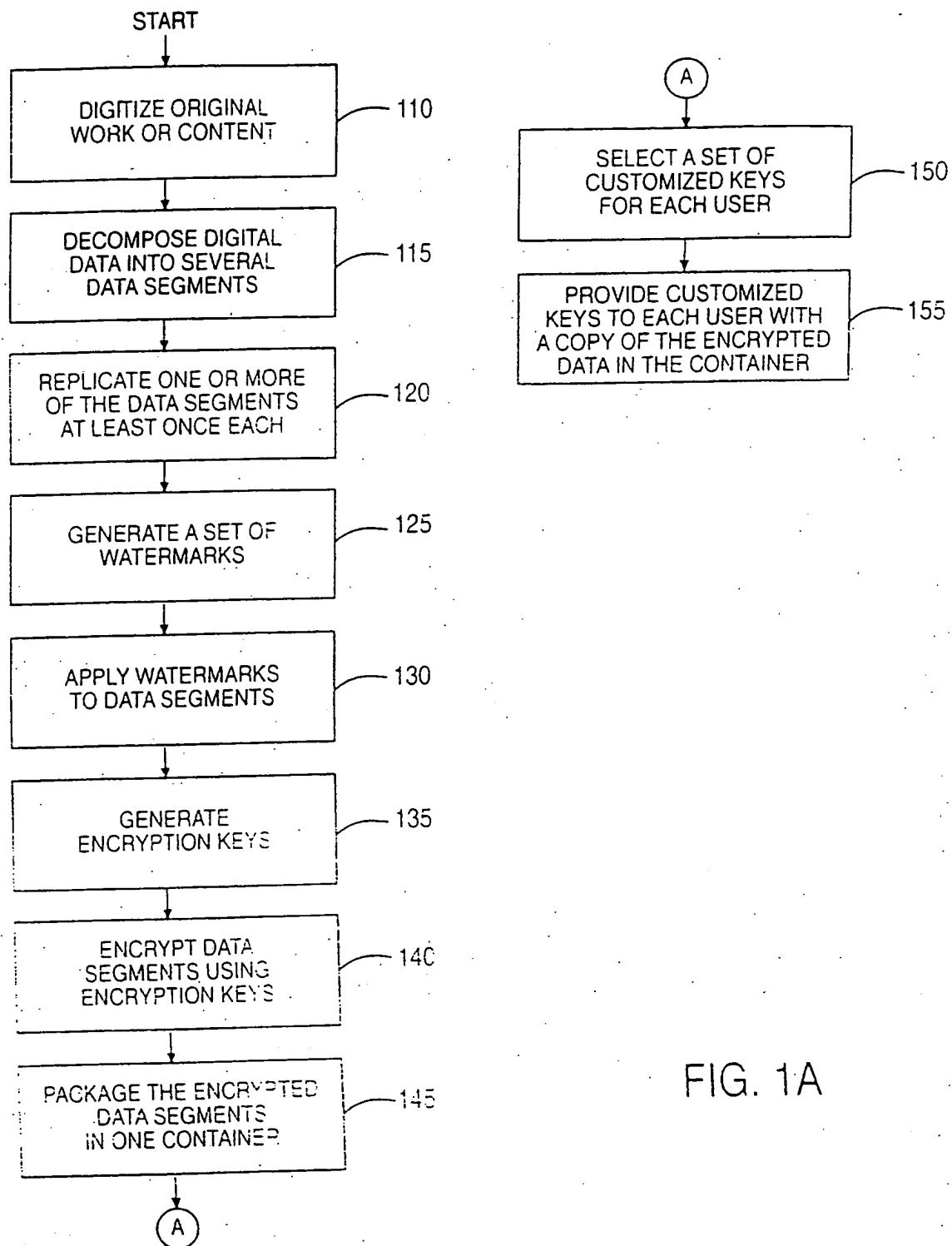


FIG. 1A

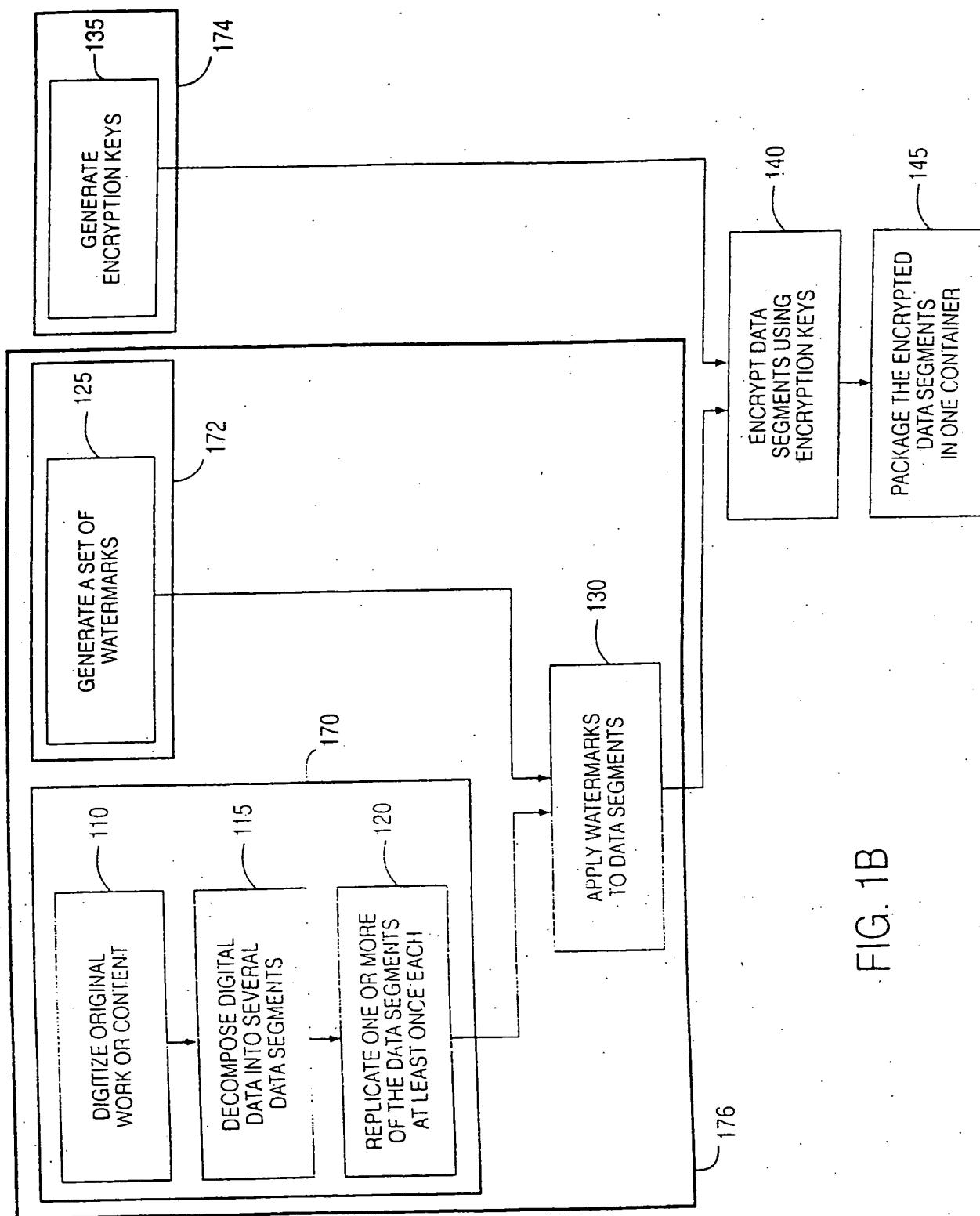


FIG. 1B

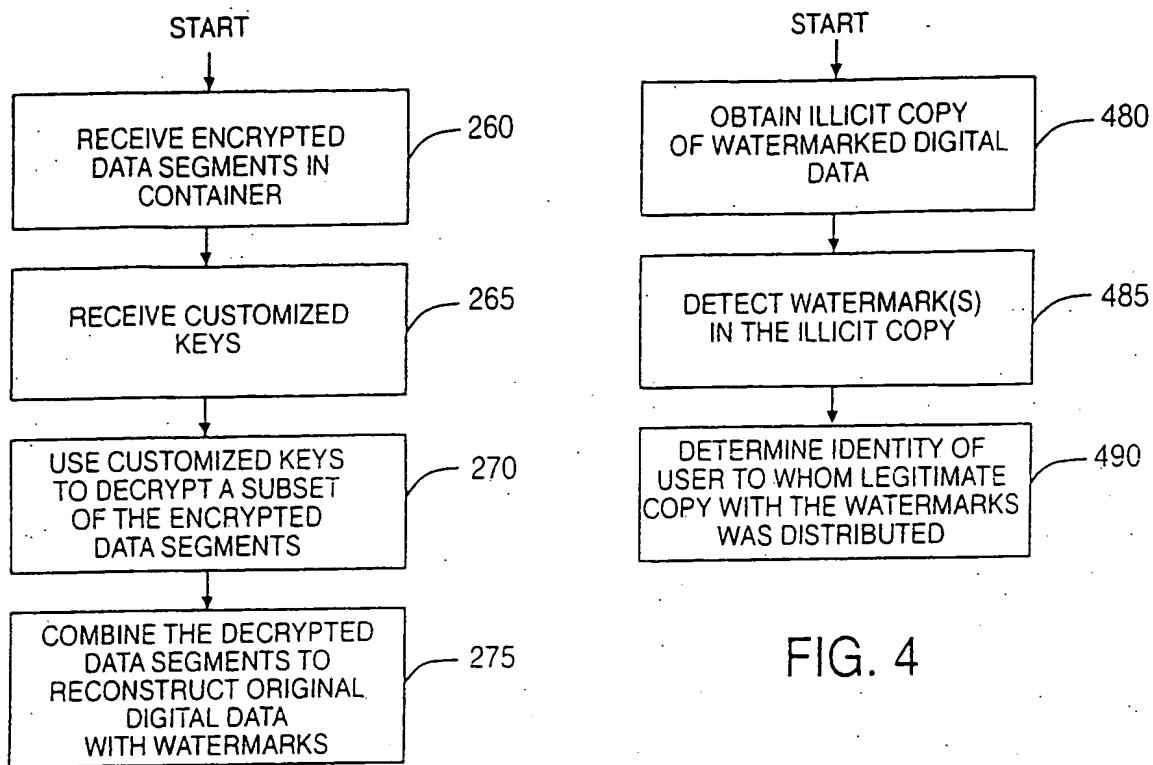


FIG. 2A

FIG. 4

4/9

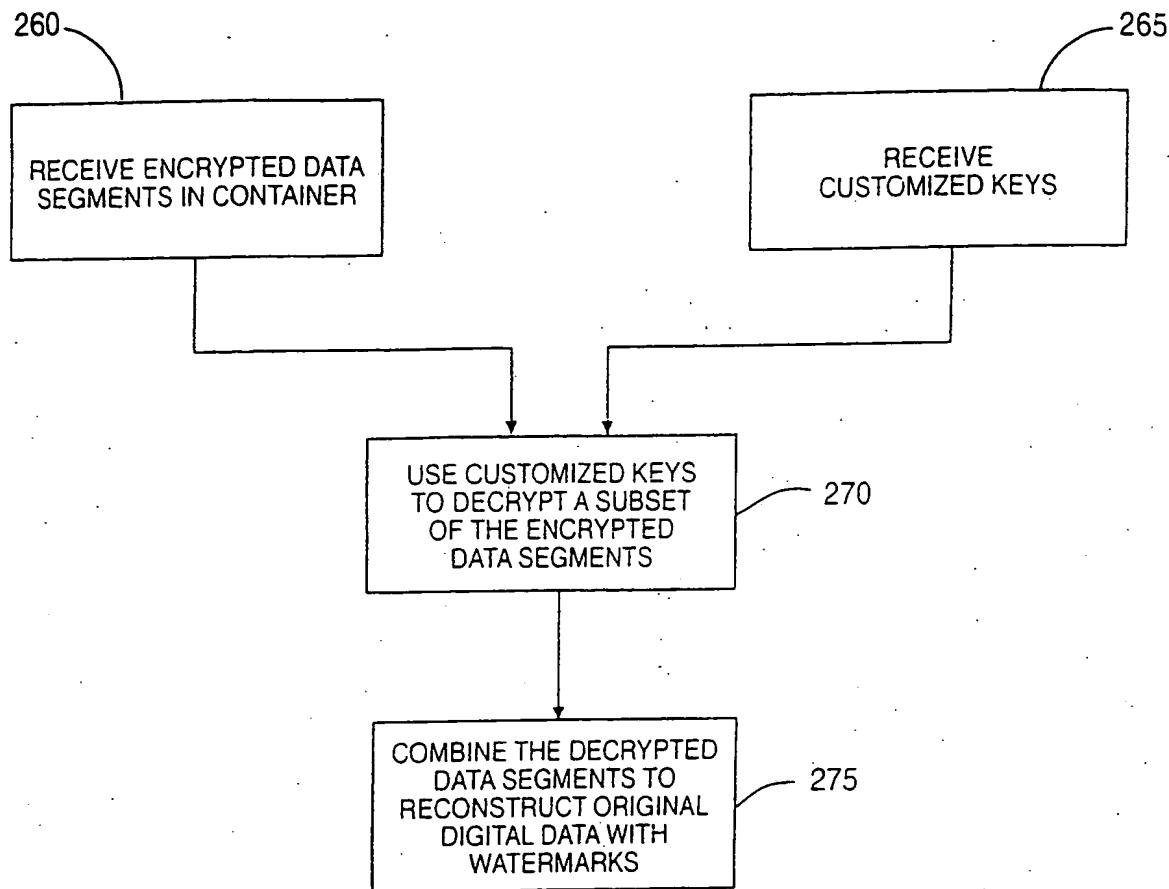


FIG. 2B

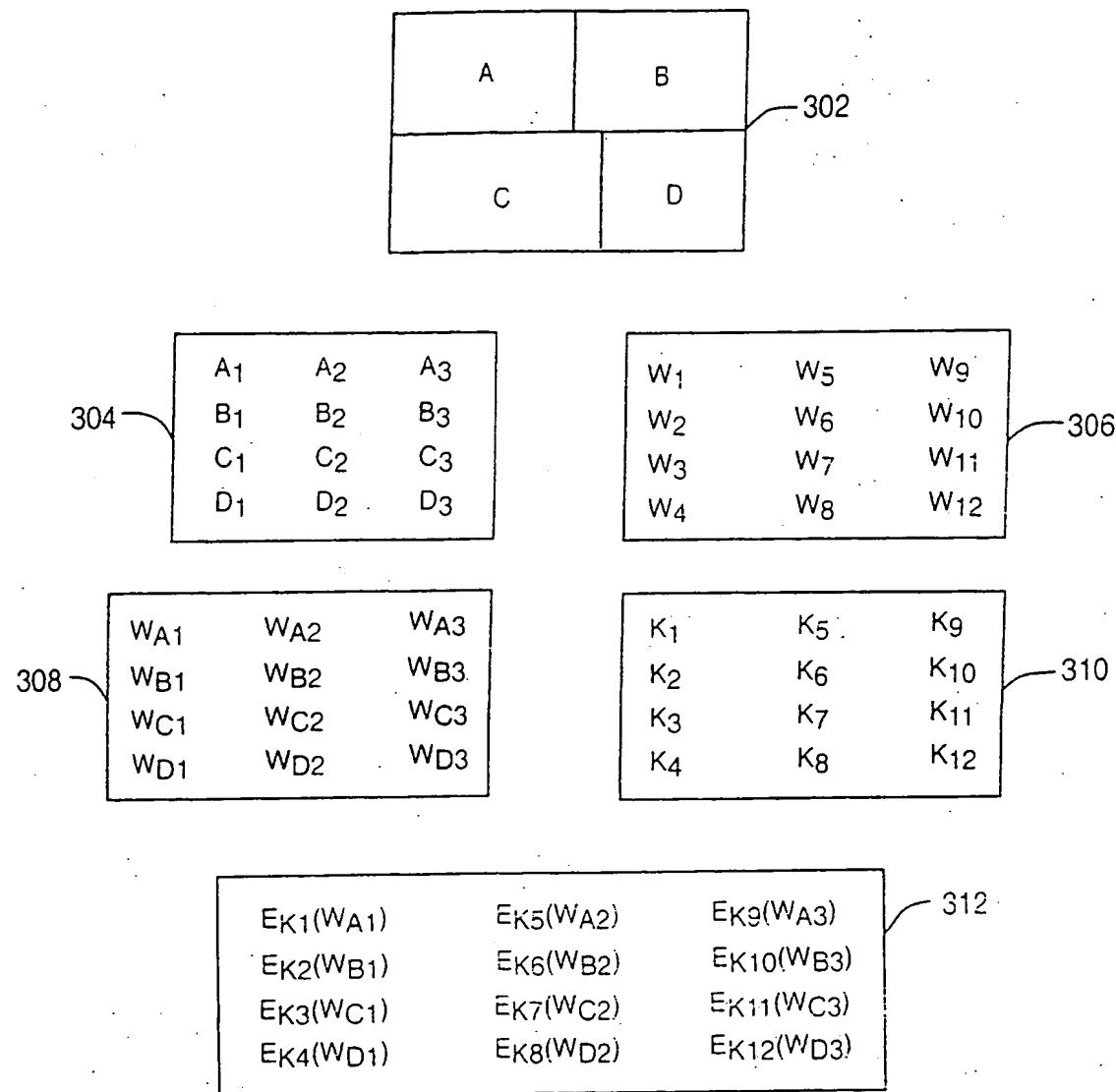


FIG. 3A

6/9

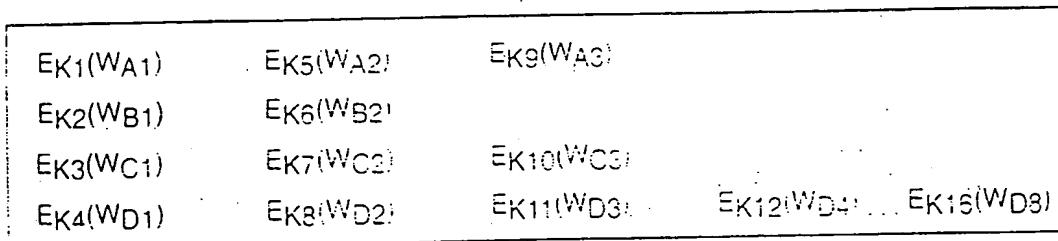
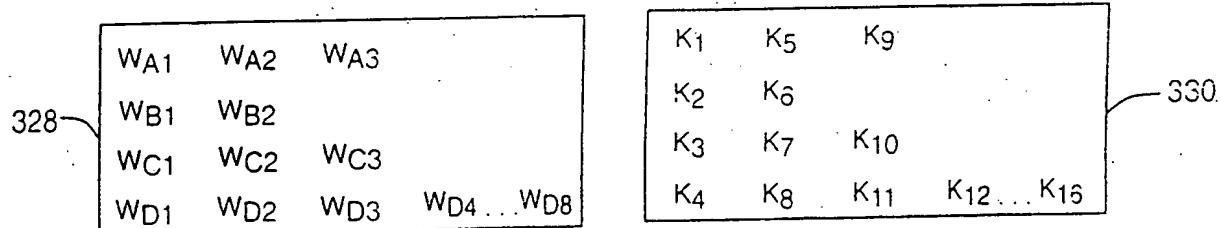
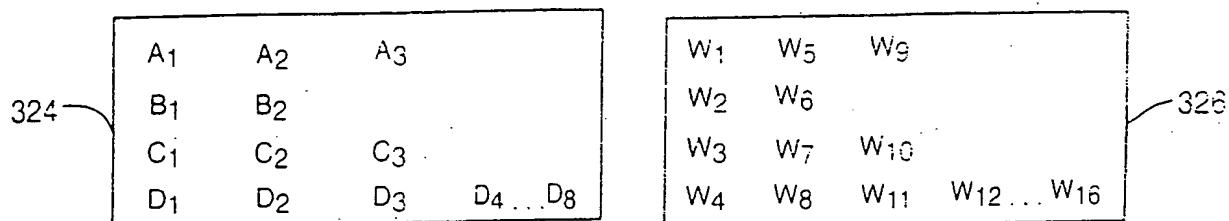
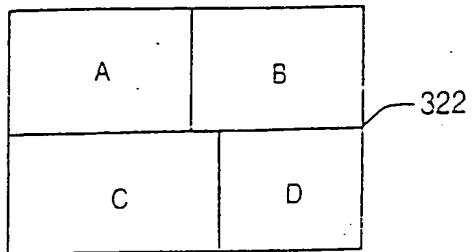


FIG. 3B

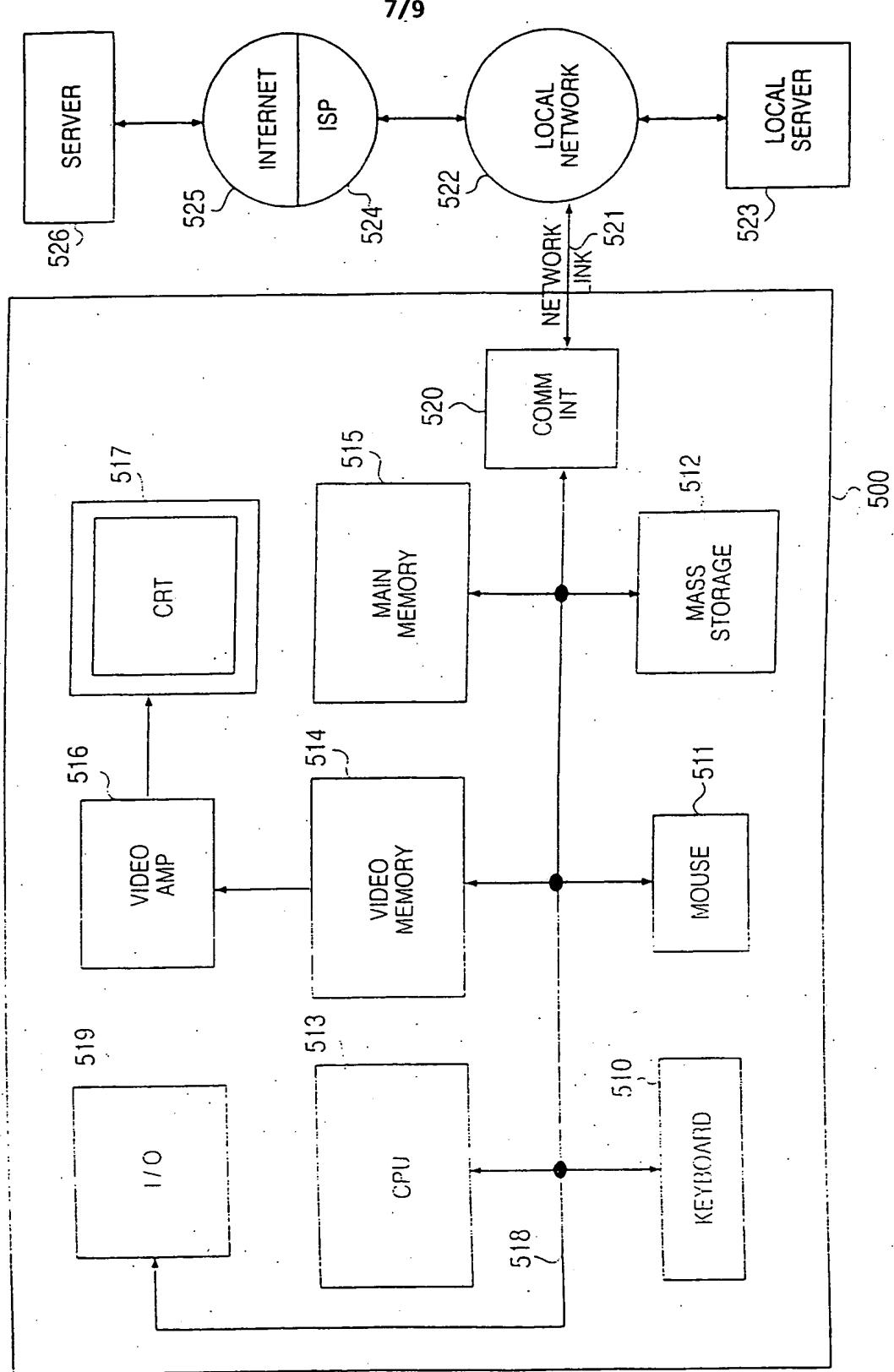


FIG. 5

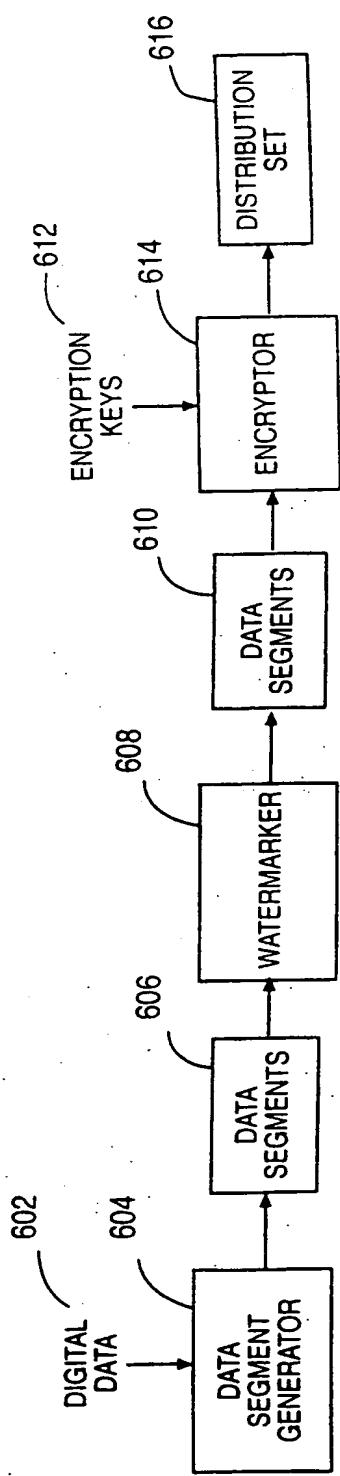


FIG. 6A

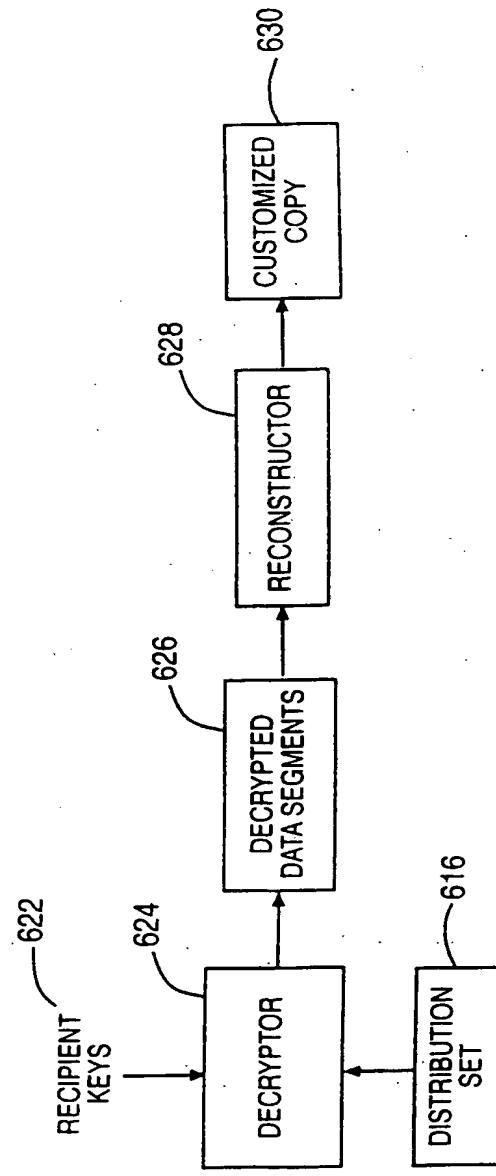


FIG. 6B

9/9

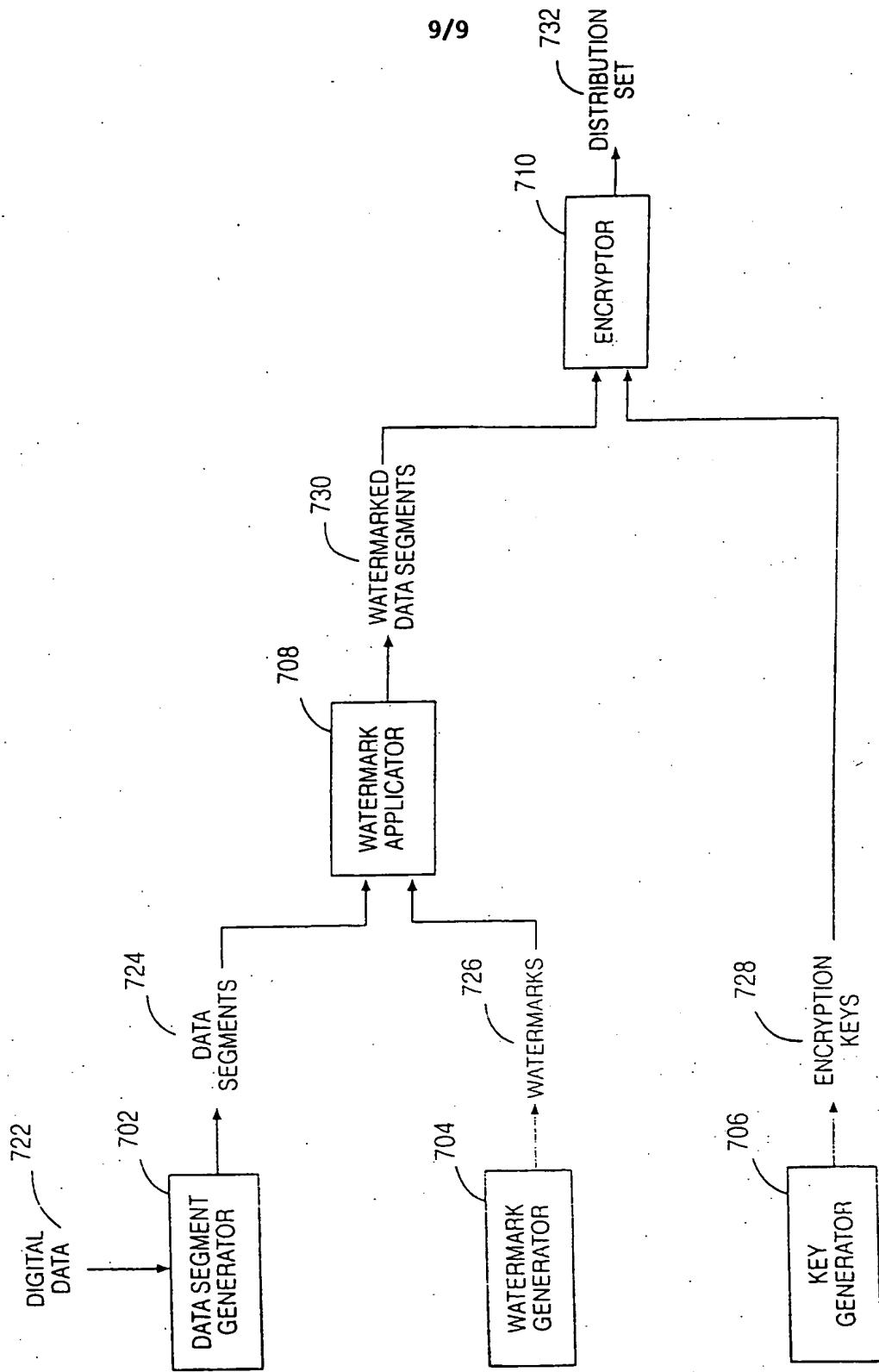


FIG. 7

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 June 2001 (21.06.2001)

PCT

(10) International Publication Number
WO 01/45410 A3

(51) International Patent Classification⁷: H04N 1/32

(21) International Application Number: PCT/US00/33151

(22) International Filing Date: 6 December 2000 (06.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/461,259 15 December 1999 (15.12.1999) US

(71) Applicant: SUN MICROSYSTEMS, INC. [US/US];
901 San Antonio Road, M/S: UPAL01-521, Palo Alto, CA
94303 (US).

(72) Inventors: CARONNI, Germano; 1063 Morse Avenue #25-300, Sunnyvale, CA 94089 (US). SCHUBA,

Christoph: 473 Hope Street #1, Mountain View, CA 94041 (US).

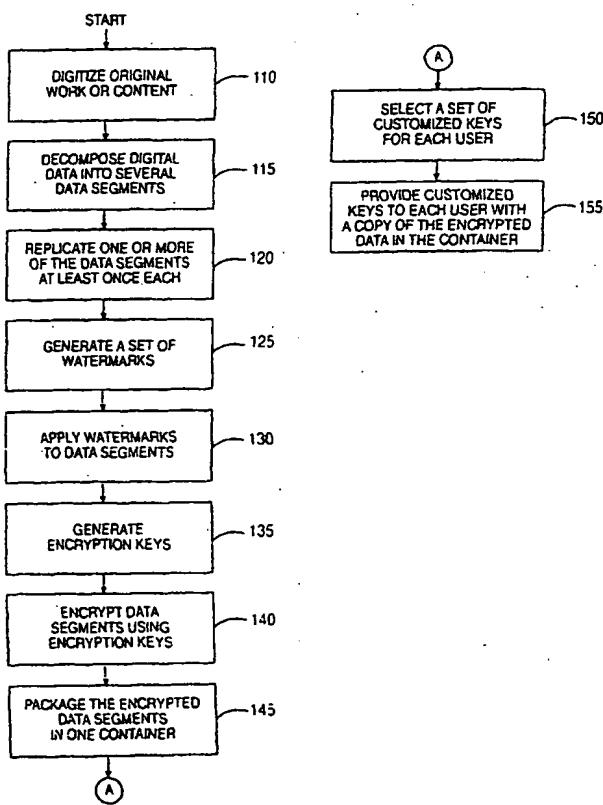
(74) Agents: HECKER, Gary, A. et al.; The Hecker Law Group, Suite 2300, 1925 Century Park East, Los Angeles, CA 90067 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,

[Continued on next page]

(54) Title: A METHOD AND APPARATUS FOR WATERMARKING DIGITAL CONTENT



(57) Abstract: A method and apparatus for watermarking digital data is described herein whereby the digital data is decomposed into a plurality of original data segments, one or more of the original data segments replicated at least once to generate replica data segments, a set of watermarks is generated, each watermark is applied to a respective data segment to generate watermarked data segments, the data segments are encrypted using encryption keys to generate encrypted data segments. One or more embodiments of the invention include providing a subset of the encryption keys corresponding to a subset of the encrypted data segments, wherein each encrypted data segment in the subset of the encrypted data segments, can be decrypted using a corresponding encryption key in the subset of encryption keys, and wherein the decrypted data segments can be combined to reconstruct the digital data including one or more of the watermarks.

WO 01/45410 A3

WO 01/45410 A3



IT, LU, MC, NL, PT, SE, TR), OAPI-patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(88) Date of publication of the international search report:
27 December 2001

Published:

— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

National Application No
PCT/US 00/33151

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 H04N1/32

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 H04N G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 99 41900 A (FRAUNHOFER GES FORSCHUNG) 19 August 1999 (1999-08-19)	1, 4, 5, 9-11, 26, 29-32, 34-36, 45, 49, 50
A	the whole document ---	13, 38
Y	EP 0 840 513 A (NIPPON ELECTRIC CO) 6 May 1998 (1998-05-06)	1, 4, 5, 9-11, 26, 29-32, 34-36, 45, 49, 50
A	the whole document ---	13, 38
		-/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search

16 July 2001

Date of mailing of the international search report

20/07/2001

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

Hazel, J

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 00/33151	
---	--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 365 589 A (GUTOWITZ HOWARD A) 15 November 1994 (1994-11-15) abstract ----	1-3, 8, 13, 15, 16, 26-28, 33, 38, 40, 41, 45
A	BRASSIL J ET AL: "ELECTRONIC MARKING AND IDENTIFICATION TECHNIQUES TO DISCOURAGE DOCUMENT COPYING" TORONTO, JUNE 12 - 16, 1994, LOS ALAMITOS, IEEE COMP. SOC. PRESS, US, 12 June 1994 (1994-06-12), pages 1278-1287, XP000496591 ISBN: 0-8186-5572-0 ----	
A	US 5 629 770 A (O'GORMAN LAWRENCE P ET AL) 13 May 1997 (1997-05-13) ----	
A	US 5 568 550 A (UR SHMUEL) 22 October 1996 (1996-10-22) ----	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/33151

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9941900 A	19-08-1999	US 6141753 A		31-10-2000
		EP 1055321 A		29-11-2000
EP 0840513 A	06-05-1998	US 5915027 A		22-06-1999
		AU 721462 B		06-07-2000
		AU 4434097 A		07-05-1998
		CA 2219205 A		05-05-1998
		JP 10145757 A		29-05-1998
		SG 63773 A		30-03-1999
US 5365589 A	15-11-1994	NONE		
US 5629770 A	13-05-1997	US 6086706 A		11-07-2000
		CA 2136166 A		21-06-1995
		DE 69421255 D		25-11-1999
		DE 69421255 T		31-05-2000
		EP 0660275 A		28-06-1995
		JP 3136061 B		19-02-2001
		JP 7222000 A		18-08-1995
US 5568550 A	22-10-1996	US 6072871 A		06-06-2000